Microsoft OpenAI

November 9, 2020

Uploaded to https://www.regulations.gov to BIS-2020-0029
With courtesy copy to Tongele.Tongele@bis.doc.gov

Mr. Tongele Tongele
c/o Office of Nonproliferation and Treaty Compliance
Bureau of Industry and Security
United States Department of Commerce
Room 2099B
1401 Constitution Avenue, N.W.
Washington, D.C. 20230

Subject: Microsoft and Open AI Comment on Advance Notice of Proposed Rulemaking (ANPRM) for the Identification and Review of Controls for Certain Foundational Technologies

References: 85 Fed. Reg. 52934 (Aug. 27, 2020) and 85 Fed. Reg. 64078 (Oct. 9, 2020); RIN 0694-AH80; Docket # 200824-0224

Dear Mr. Tongele:

Microsoft and OpenAI appreciate the opportunity to comment on the Advance Notice of Proposed Rulemaking (ANPRM) for the Identification and Review of Controls for Certain Foundational Technologies. We support the Department of Commerce's efforts under the Export Control Reform Act of 2018 (ECRA) to evaluate the appropriate nature and scope of export restrictions on the most important digital technologies, together with industry, academics, and others. We recognize that US national security concerns are at the heart of these efforts, and we share Commerce's desire that any restrictions enhance rather than undermine US national security. To achieve this, targeted controls focused on end users and uses of concern are needed that protect against national security risks on the one hand, while preserving the beneficial uses and US technological leadership on the other.

To make these controls more effective and dynamic, we propose their digital transformation – a new approach that would deploy novel digital solutions within the technologies themselves. These solutions would directly enforce and monitor government-imposed controls on users and uses, and secure the infrastructure surrounding the technologies to decrease the risk that controls will be subverted. Key features include:

- **Software features** designed into sensitive technologies to enable real-time controls against prohibited uses and users. These features would include, for example, identity verification systems and information flow controls to discern whether facts and criteria are consistent with authorized users and uses. "Tagging" can be used to ensure the same controls apply to derivatives of these sensitive technologies.

- **"Hardware roots of trust"** built into hardware that contains sensitive technologies can complement software-based solutions by requiring authorization for access. More robust **hardware identity verification** through secure co-processors akin to those used, for example, to secure payment in mobile phones or to prohibit in-game cheating in game consoles can further protect hardware against unauthorized access and uses.

- **Tamper-resistant tools** for sensitive technologies and for protective software and hardware solutions themselves to harden infrastructures against subversion.

- At a minimum, the above techniques can enhance export controls. **Artificial intelligence techniques**, however, can be used to more adeptly identify and restrict problematic end users or uses, including through continuous improvement and learning.

These solutions should be reserved for the most sensitive technologies and be employed transparently. Employed appropriately, however, they can provide a far more powerful, dynamic, and targeted method for controlling exports of these important technologies. Multilateral coordination will also be vital to preserving beneficial uses of these technologies and US technological leadership.

A. End User and End Use-Based Controls Are the Best Approach for Export Controls on Foundational Technologies

1. End User and End Use-Based Controls Are the Best Approach for Export Controls on Emerging Technologies

In response to the ANPRM regarding Review of Controls for Certain Emerging Technologies, Microsoft, OpenAI and many others noted the challenges of imposing export controls on the technologies themselves, rather than on particularly problematic uses and users.[1] Importantly, technology with beneficial and problematic uses cannot be distinguished based on performance capabilities or other

[1] *See, e.g.* Microsoft Comment on Advance Notice of Proposed Rulemaking Regarding Review of Controls for Certain Emerging Technologies ("Microsoft Comment"), at 3, *available at* https://www.regulations.gov/document?D=BIS-2018-0024-0175; OpenAI Comment on Advance Notice of Proposed Rulemaking Regarding Review of Controls for Certain Emerging Technologies ("OpenAI Comment,") *available at* https://www.regulations.gov/document?D=BIS-2018-0024-0195; Google Comment on Advance Notice of Propose Rulemaking Regarding Review of Controls for Certain Emerging Technologies, at 19, *available at* https://www.regulations.gov/document?D=BIS-2018-0024-0160; Semiconductor Industry Association Comment on Advance Notice of Proposed Rulemaking Regarding Review of Controls for Certain Emerging Technologies, *available at* https://www.regulations.gov/document?D=BIS-2018-0024-0130.

.

technical criteria. The very same technology can be used both as a powerful tool and a powerful weapon. Responses also highlighted the substantial progress companies in foreign countries have made towards the development of emerging technologies, allowing these companies to fill any void created by export restrictions on US companies. They pointed out the importance to US industry of access to global markets and global talent, the ability to sell products that can be used globally, and the risk that – if the US restricts this access and ability – companies based in countries not bound by such restrictions will overtake US companies in technological development. If the most important emerging technology is developed abroad, they emphasized, those who are not invested in US interests and priorities, including those who may be hostile to them, will have access to more powerful technology than the US will – an outcome that undermines the core goals of the US export control system. In addition, export controls placed directly on emerging technologies based on their performance capabilities are ill-fitted to the rapid pace of development in these areas and are likely to be outdated almost as soon as they are implemented.

The importance of all these problems is heightened for any emerging technologies under consideration for "dual use" controls because they are generally not technologies that had their origin in the military or intelligence sectors and then became useful for a limited set of commercial or other beneficial purposes. Instead, these technologies encompass broad-based capabilities like computing power and artificial intelligence grown out of the commercial and consumer sector, whose beneficial applications outnumber any problematic ones. For this reason, Microsoft's Emerging Technologies ANPRM comments urged Commerce to consider such technologies "common use" rather than simply "dual use," and stressed the importance of controls only on their problematic end uses and end users.

2. For Many of the Same Reasons as with Emerging Technologies, End User and End Use-Based Controls Are the Best Approach for Export Controls on Foundational Technologies

The same approach is appropriate for foundational technologies. ECRA does not precisely define foundational technologies, except to state that these should be technologies essential to the national security interests of the United States. At a minimum, foundational technologies should also be those technologies required for the design, development, production, or use of emerging technologies that are essential to those interests, as the statute treats emerging and foundational technologies together. The statute also excludes from consideration as foundational technologies those technologies already subject to control. This is because Congress was concerned with both emerging and foundational technologies that are outside the scope of current list-based controls because they are not tied to the design, development, production or use of a single, specific article on the Commodity Control List, but instead relate to many articles and many uses. These technologies tend to be information or digital technologies. With these technologies – whether emerging or foundational – only by focusing on *who* is using them and for *what* purpose, can Commerce stop their problematic uses without undermining US leadership in them and their many beneficial uses.

Like with emerging technologies, it is not possible to make any meaningful distinction between beneficial versus nefarious uses of technologies potentially under consideration for foundational controls *based only on its performance capabilities or other technical criteria*. Using facial recognition as an example, the same digital biometrics technology, as well as software and hardware, capture and analyze the information, regardless of ultimate use.[2] The same camera or digital voice recorder, the same data storage and processing computers, and the same software algorithms can be used to digitally identify people – whatever the purpose. The amount of computing power used and the acuity of the recognition algorithms employed are the same whether one is scanning a crowd to find a missing child, a criminal, or a terrorist, versus to find an oppressed dissident or minority. The positive or negative impact from the technology is instead the result of who applies the system and to what end.

Also like with emerging technologies, technologies potentially under consideration for foundational controls are currently developed and enhanced around the world, both by foreign companies and by foreign workers supporting US companies. Artificial intelligence, for instance, is being broadly developed worldwide. So are facial recognition capabilities.[3] Access to global R&D, talent, and markets is no less important for foundational technologies than it is for emerging. Efforts to restrict the technology itself will therefore not be effective, will only constrain US competitiveness and threaten its technological leadership, and will risk placing the most important foundational technologies in the hands of potential adversaries.

Moreover, like with emerging, technologies potentially under consideration as foundational are constantly evolving and improving, such that any controls based on their performance criteria would become quickly outdated. Facial recognition algorithms, for example, have improved in accuracy at a remarkable rate in recent years, due to an increased reliance on deep neural networks. More recently, the same pattern has begun to play out for AI systems that can search, classify, and generate text information.

As such, foundational technologies are also best regulated by restrictions on the end users that can have access to them, and the end uses to which they can be put.

B. Commerce Should Consider A Digital Transformation of End User and End Use-Based Controls

Through all the challenges posed by foundational technologies, we see an opportunity. End user and end use export restrictions for modern technologies can be made better through a digital transformation of these controls.[4] At the core of this approach is that digital solutions incorporated into the technologies themselves can more effectively and flexibly control the technologies in a way that advances US interests rather than inhibits them. These solutions can implement and enforce

[2] By referring to facial recognition, AI, or any other technology in this comment, Microsoft and OpenAI do not intend to suggest that those technologies would constitute "foundational" technologies for purposes of ECRA. They are intended as examples for purposes of discussion only.

[3] https://www.biometricupdate.com/201909/yitu-and-visionlabs-impress-in-latest-nist-facial-recognition-test-results (top performers mentioned include Paravision (US); Hikvision and Yitu (China); and Vision Labs (Netherlands)).

[4] Though we are proposing this approach in response to the ANPRM on foundational technologies, this approach would equally make controls on emerging technologies more effective.

government-created restrictions on inappropriate uses and users, as well as secure the infrastructure surrounding these technologies to prevent subversion. They can continuously adapt and update in response to new information or restrictions, allowing for a more nimble and responsive export controls approach.

We understand that this proposed digital transformation would constitute a significant change in how Commerce approaches export controls, but we strongly believe it is necessary to effectively navigate the promise and risk of the most important technologies in the US and worldwide. The world in which traditional compliance and enforcement techniques can keep technologies out of the hands of our adversaries will soon be entirely behind us. While we acknowledge the effort required to create and monitor these tools, ultimately a digitally transformed export control system – in addition to being more effective – would be less burdensome than the traditional compliance and enforcement techniques of today.

1. The Basic Structure of Technological End User and End Use-Based Controls

Digitally transformed controls will have several components that work together to implement and enforce end user and use-based restrictions imposed by the government. Software and hardware-based solutions can vet identity, control information flow, and authorize and deny access based on important information about users and use. These components can be enforced and assured by a secure trusted computing base. Tamper-resistant solutions can harden infrastructure and prevent circumvention of controls. As a simpler, less robust, but familiar example, these tools are akin to mechanisms that platforms use to determine which applications are placed in application stores, and which (because of concerns like privacy, security, or reliability) must be denied.

User and Use Verification and Control

Identity verification systems can be used to determine who is authorized to use certain technology. These can be paired with secure co-processors that provide a robust cryptographic machine identity. As an example of how this works, most mobile phones today include a secure co-processor with a hardware identity to facilitate secure payment. Gaming consoles have long used security co-processors to secure the console itself against unauthorized modification, to protect game authors' creations from unauthorized duplication, and to prohibit in-game cheating. For particularly sensitive technologies, secure co-processors can be integrated into broader systems, where devices will only work with other devices whose identity and security status can be verified.

In addition, information flow controls can prevent outputs from going to unauthorized users. Digital rights management is one familiar information flow control, but there are many others. Outputs can be required to remain on the source platform unless designated criteria are met, such as up-to-date security or a specified set of users or a physical location. This process can be controlled by access control managers embedded within the technology itself. Particularly sensitive data can also be "tagged," and systems can be set up to restrict the flow of the tagged data unless and until it undergoes an affirmative un-tagging process.[5]

The flow of information within a technology system – who can see what information – is also key to understanding how it is used. To give a commonly understood example, malicious applications on mobile phones can be identified through controls designed to flag when an application is accessing a user's contacts or the contents of text messages.

Tamper-proof Hardware and Secure Infrastructure

Security co-processors, access control managers, and other associated machine identity tools can also be used to protect hardware against tampering. When a component is manufactured, a manufacturer can attest to its identity and these attestations can be shipped with the component. As that component is included into larger components or integrated systems, the identity of the larger system can include the identities of its components.

Such hardened systems will not only help ensure that technologically enabled export controls cannot be subverted by adversaries, they can be used to greatly enhance security across critical infrastructure and supply chains with important benefits for commercial, privacy, intelligence, and other interests.

AI-enabled Identification of Problematic Uses and Users

At a minimum, the techniques described above can be used to enforce end use and end user-based export controls issued by the government. Artificial intelligence coupled with these techniques, however, can even more dynamically help identify and block problematic users or uses.

Such AI solutions are already in development. Consider OpenAI's GPT-3, a large neural language model trained on a broad range of internet data. Given the right "prompt," GPT-3 can compose stories, answer questions, write poetry, have a dialogue with the user, write programs in Javascript or Python, and perform many other tasks. GPT-3 itself, however, needs to be controlled in order to prevent it from potentially performing problematic tasks, such as providing a user instructions on how to build an explosive device or exploit a security vulnerability, creating malicious code, or generating racist, sexist, or otherwise unacceptable content. Technological controls that OpenAI is developing for this purpose can similarly be used as a component of digital solutions for end user and use-based export controls.

The technological controls OpenAI is working to implement for GPT-3 are multi-faceted, including such measures as i) human raters who give feedback that is then used to train the model, ii) inspections of representations within the model to determine and block characteristics of undesirable interactions, and iii) use of a specially trained copy of GPT-3 itself to identify and flag problematic interactions. In essence, the AI can be trained to mimic the nuanced judgement of human experts and can do so continuously in the deployed environment rather than just once at the point of initial export. We expect

[5] These mechanisms can also provide a durable record of information flow and device and user access that can be used, even after the fact, to audit end user access for the purposes of enforcement actions or making future enhancements to lists of restricted users or to the systems themselves.

that, over time, further AI techniques can be used to identify 'unusual' uses on a per-user basis, then take appropriate action (e.g., restrict application for a particular use, route an unexpected usage pattern for further human review).

Moreover, AI-based systems can improve over time in ways that export controls today cannot. For example, they can be retrained or redeployed based on observed authorized and unauthorized use attempts. This could help Commerce learn more about potentially problematic uses and users to inform adjustments to export controls. This would require an evolution in regulatory approach, but it would ultimately give Commerce even greater tools to accomplish its important ends.

2. US Industry Is Already Working To Incorporate Technological End User and End Use-Based Controls into Powerful Technologies

US industry has its own imperatives – separate from export controls – to ensure that its most important technologies are not used in destructive and dangerous ways. These imperatives help ensure that such a digital transformation of controls will be successful, and also help enhance US technological leadership.

These imperatives come, in part, from companies' corporate social responsibility commitments. For example, Microsoft has long publicly supported governmental restrictions on the use of facial recognition technology, and has committed to self-enforce similar restrictions based on its Facial Recognition Principles.[6] More recently, Microsoft has imposed gating restrictions on its Custom Neural Voice service, a technology that creates a synthetic voice based on audio data from real speakers. This technology has incredible benefits, such as allowing people with degenerative diseases to preserve their own voices to project from a computing device when they can no longer speak. Because the technology can also be used to create deep fakes, however, Microsoft restricts access to the technology based on use and users.[7] Technology can also identify non-conforming uses and users, for instance, by matching the voice of the potential user of the service to the audio files from which he or she wishes to have a synthetic voice created.

Business interests also drive industry interest in effective end user and end use controls. Customers want certain assurances from companies about their technologies. In the GPT-3 example above, not only are the problematic uses described contrary to OpenAI's mission – "to ensure that artificial general intelligence (AGI)—by which we mean highly autonomous systems that outperform humans at most economically valuable work—benefits all of humanity"[8] – it is unlikely a commercial or consumer application of GPT-3 could survive without controls to prevent these uses. Customers who want to deploy technologies like GPT-3 have expectations of reliability and safety. OpenAI and its customers are already collaborating on AI-driven systems to edit model outputs so that they conform to customer

[6] https://blogs.microsoft.com/on-the-issues/2018/07/13/facial-recognition-technology-the-need-for-public-regulation-and-corporate-responsibility/

[7] https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/concepts-gating-overview

[8] https://openai.com/about/

expectations, like maintaining a consistent tone of voice or being able to provide reliable safeguards against user-generated hate speech.

Similarly, Microsoft is deeply involved in the Open Compute project, a collaborative community focused on hardware technology. One of the core objectives of Open Compute is hardware security, i.e., providing technology companies with openly available tools needed to secure systems, even where those systems involve hardware from multiple companies. Open Compute security projects include those focused on secure boot and boot code integrity that ensure hardware only does what it is intended to do, as well as attestations of origin of each element of a system, including in complex supply chains.[9] Microsoft incorporates chips that use Open Commute Project-consistent standards into its Azure systems.

Companies also have a direct commercial interest in using such controls to protect customer information from inappropriate and unwelcome uses. Companies' ability to protect both legal and voluntary customer privacy commitments would be greatly enhanced by tools that can track and control the flow of information within systems, and to secure against unwanted flows out of those systems. Where customer data is at risk of being shared with US adversaries, this privacy interest also becomes a US national security one.

3. Our Proposed Approach Helps Advance the Broader US National Security Strategy

The White House's recently released National Strategy for Critical and Emerging Technologies report further underscores the importance of the approach we are proposing.[10]

The report emphasizes what is critical to US national security is that the US lead in high priority technology areas. As we discussed above, placing more restrictive than necessary impediments on US development of foundational technology would jeopardize this leadership. The report notes that maintaining US leadership is instead accomplished, first and foremost, by ensuring a skilled workforce, increasing private investment in technology, decreasing regulatory obstacles, and increasing government investment in research and development.[11]

The report also recognizes that the US must protect the technological advantages that result from its leadership. While export controls are mentioned as a tool to help ensure this protection, the report's primary focus is on building security into technology itself, increasing security in research institutions, and securing the supply chain.[12] These goals would be enhanced by the approach we propose above.

[9] https://www.opencompute.org/projects/security; https://www.opencompute.org/wiki/Security

[10] https://www.whitehouse.gov/wp-content/uploads/2020/10/National-Strategy-for-CET.pdf

[11] *Id.* at 7-8.

[12] *Id. at* 9-10.

Conclusion

For all of these reasons, we believe it is time for the digital transformation we have proposed. US government and US commercial interests are aligned in the need for secure technological control environments that can enable the wide range of beneficial uses for sensitive and important technologies, while protecting against improper, even dangerous, ones.

We look forward to working with Commerce to begin this transformation. If you have additional questions or would like to discuss the comments further, please contact Sarah O'Hare O'Neal at Sarah.ONeal@microsoft.com or (202) 365-9011 or Jack Clark at jack@openai.com or (415) 685-1845.

/s/
Sarah O'Hare O'Neal
Partner, Associate General Counsel,
Global Trade
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052

/s/
Jack Clark
Policy Director
OpenAI
3180 18th Street
San Francisco, CA 94110